

CLAIMS

We claim:

- Sub A2 5
1. A computer-readable medium having computer-executable instructions for communicating between an application and a multipoint processing module having at least one audio processor module for processing audio data in a multipoint conference and at least one video processor module for processing video data in a multipoint conference, the computer-executable instructions performing the step of:

exposing at least one interface by the multipoint processing module to

receive a request from the application to command the multipoint processing module to modify its default operation to alter at least one attribute of at least one of the audio processor module and video processor module.

2. The computer-readable medium of claim 1 wherein said at least one interface comprises an audio interface, the application using said audio interface to request the multipoint processing module to change a routing of at least one audio input stream towards at least one audio output stream.

3. The computer-readable medium of claim 2 wherein the request is selected from the group consisting of:

a command to retrieve an audio crossbar topology, the audio crossbar topology indicating how a set of audio input streams is being routed to a set of audio output streams;

a command to change the audio crossbar topology to indicate to the multipoint processing module how the set of audio input streams should be routed to a set of audio output streams;

a command to retrieve a value of an audio crossbar control parameter;

5 a command to set a value of an audio crossbar control parameter;

a command to retrieve a minimum value, a maximum value, and a default value for an audio crossbar control parameter;

a command to retrieve a mixing capability and a transcoding capability of the audio crossbar; and

10 a command to retrieve an audio level of a list of audio input streams.

4. The computer-readable medium of claim 1 wherein said at least one interface comprises a video interface, the application using said video interface to request the multipoint processing module to change a routing of at least one video input stream towards at least one video output stream.

5. The computer-readable medium of claim 4 wherein the request is selected from the group consisting of:

a command to retrieve a video crossbar topology, the video crossbar topology indicating how a set of video input streams is being routed to a set of video output streams based on a content of associated audio input streams;

20 a command to change the video crossbar topology to indicate to the multipoint processing module how the set of video input streams should be routed

to a set of video output streams based on a content of associated audio input streams;

a command to retrieve a value of a video crossbar control parameter;

a command to set a value of a video crossbar control parameter;

5 a command to retrieve a minimum value, a maximum value, and a default value for a video crossbar control parameter;

a command to retrieve a mixing capability and a transcoding capability of the video crossbar; and

a command to retrieve a video level of a list of video input streams.

10

6. The computer-readable medium of claim 2 wherein said at least one interface further comprises a video interface, the application using said video interface to request the multipoint processing module to change a routing of at least one video input stream towards at least one video output stream.

15

7. The computer-readable medium of claim 6 wherein

the request to route at least one audio input stream is selected from the group consisting of:

a command to retrieve an audio crossbar topology, the audio

20

crossbar topology indicating how a set of audio input streams is being routed to a set of audio output streams;

a command to change the audio crossbar topology to indicate to the multipoint processing module how the set of audio input streams should be routed to a set of audio output streams;

5 a command to retrieve a value of an audio crossbar control parameter;

a command to set a value of an audio crossbar control parameter;

a command to retrieve a minimum value, a maximum value, and a default value for an audio crossbar control parameter;

10 a command to retrieve a mixing capability and a transcoding capability of the audio crossbar; and

a command to retrieve an audio level of a list of audio input streams;

the request to route at least one video input stream is selected from the group consisting of:

15 a command to retrieve a video crossbar topology, the video crossbar topology indicating how a set of video input streams is being routed to a set of video output streams based on a content of associated audio input streams;

20 a command to change the video crossbar topology to indicate to the multipoint processing module how the set of video input streams should be routed to a set of video output streams based on a content of associated audio input streams;

a command to retrieve a value of a video crossbar control parameter;

a command to set a value of a video crossbar control parameter;

a command to retrieve a minimum value, a maximum value, and a default value for a video crossbar control parameter;

a command to retrieve a mixing capability and a transcoding capability of the video crossbar; and

a command to retrieve a video level of a list of video input streams.

8. The method of claim 7 wherein said at least one interface further comprises a format control interface, the application using said format control interface to retrieve and set an audio format and a video format, the format control interface comprising:

a command to retrieve a preferred audio and video format for a conference;

a command to set the preferred audio and video format for the conference;

a command to retrieve a format structure and configuration capability structure pair of a conference, the format structure and configuration capability structure pair describing an audio and video format supported by the conference;

a command to retrieve a number of audio and video format structure and configuration capability structure pairs that are available in a conference;

a command to reorder a list of preferred audio formats; and
a command to reorder a list of preferred video formats

- 5 9. The computer-readable medium of claim 3 wherein the audio crossbar control parameter is selected from a group of audio crossbar control parameters, the group comprising:

a setting to specify a periodicity of an interrupt service routine;
a setting to specify a maximum number of mixed input signals;
10 a setting to enable and disable silence detection;
a setting to enable and disable silence compression; and
a setting to enable and disable automatic gain control.

10. The computer-readable medium of claim 3 wherein the multipoint processing module
15 disables the command to set a value of an audio crossbar control parameter when a control flag is set.

11. The computer-readable medium of claim 5 wherein the video crossbar control parameter is selected from a group of video crossbar control parameters, the group
20 comprising:

a setting to specify a first time to evaluate whether a speaker is continuing
to speak;

a setting to specify a second time during which a speaker and a video switching process can not be taken over by a second speaker; and

a setting to specify a third time, the third time being the time when a switch is made and when a fast update request is sent to the speaker's system.

5

12. The computer-readable medium of claim 5 wherein the multipoint processing module disables the command to set a value of a video crossbar control parameter when a control flag is set.

10 13. A method to communicate between a media service provider component and a multipoint processing module controlling an encoder module and a decoder module for processing video data in a multipoint conference, the method comprising the step of:

15 exposing at least one interface by one of the media service provider component and the multipoint processing module to communicate commands and indications between the media service provider component and the multipoint processing module.

20 14. The method of claim 13 wherein said at least one interface further comprises a pin interface, the multipoint processing module using said pin interface to retrieve a direction and crossbar positional index of one of the audio streams and video streams.

15. The method of claim 13 wherein said at least one interface further comprises a decoder interface to handle decoder commands, the decoder interface comprising:

- a command to complete updating a video frame and display the video frame until commanded to release the video frame; and
- 5 an indication of a video temporal and spatial trade-off of the encoder.

16. The method of claim 13 wherein said at least one interface further comprises an encoder interface to send encoder commands to the encoder, the encoder interface comprising:

- 10 a command to enter a fast-update mode;
- a command to perform a fast update of a group of blocks;
- a command to perform a fast update of a macroblock;
- a command to use sync for every group of blocks;
- an indication that a set of macroblocks has been received with errors and
- 15 has been treated as not coded; and
- a command to set a relative tradeoff between a high spatial resolution and a high frame rate.

17. The method of claim 13 wherein the multipoint processing module has a video pin, said at least one interface further comprises a network statistics interface to communicate network characteristics between the video pin and to the media service provider component, the network statistics interface comprising:

- 20 a command to inform the video pin of error channel conditions;

a command to supply the media service provider component the error channel conditions;

5 a command to retrieve values of the error channel conditions with which the video pin may be setup, the values including a minimum value, a maximum value, a default value, and a support value;

a command to inform the video pin a channel packet loss rate;

a command to supply the media service provider component the channel packet loss rate; and

10 a command to retrieve values of the channel packet loss rate with which the video pin may be setup, the values including a minimum value, a maximum value, a default value, and a support value.

18. The method of claim 13 wherein the multipoint processing module has a video pin, said at least one interface further comprises a bandwidth interface comprising:

15 a command to specify an upper limit in bandwidth transmission of the video pin;

a command to retrieve the video pin's upper limit in bandwidth transmission;

20 a command to retrieve values of the upper limit in bandwidth transmission with which the video pin may be setup, the values including a minimum value, a maximum value, a default value, and a support value;

19. The method of claim 13 wherein the multipoint processing module has a video pin,
said at least one interface further comprises a frame rate control interface comprising:

a command to specify a video frame's average display time to the video
pin;

5 a command to retrieve the video frame's average display time;

a command to retrieve values for the video frame's average display time
with which the video pin may be setup, the values including a minimum value, a
maximum value, a default value, and a support value.

10 20. The method of claim 13 wherein the multipoint processing module has a video pin,
said at least one interface further comprises an RTP packet interface comprising:

a command to adjust a maximum RTP packet size generated by the video
pin;

15 a command to supply the media service provider component the maximum
RTP packet size; and

a command to retrieve values for the maximum RTP packet size with
which the video pin may be setup, the values including a minimum value, a
maximum value, a default value, and a support value.

20 21. A multipoint processing accelerator apparatus for transmitting audio and video data
over a plurality of channels in a multipoint conference being controlled by an
application, the apparatus comprising:

at least one hardware module having a default operation for applying signal processing operations to at least one of the audio and video data; and a minidriver, said minidriver communicating with the application through at least one property set to do one of receiving a command to modify the default operation of the at least one hardware module and sending a command to the application.

22. The apparatus according to claim 21 wherein the at least one property set comprises an audio topology property set.

23. The apparatus according to claim 22 wherein the audio topology property set comprises:

a property to do one of updating an audio crossbar content and retrieving an audio crossbar content;

a property to retrieve mixing and transcoding capabilities of an audio crossbar;

a property to do one of setting a periodicity of an interrupt service routine and getting a periodicity of an interrupt service routine;

a property to do one of setting a maximum number of mixed input signals and getting a maximum number of mixed input signals;

a property to do one of enabling silence detection and disabling silence detection;

a property to do one of enabling automatic gain control and disabling automatic gain control; and

a property to retrieve a value of an audio level of a list of audio input streams.

5

24. The apparatus according to claim 21 wherein the at least one property set comprises a video topology property set.

25. The apparatus according to claim 24 wherein the video topology property set comprises:

10

a property to do one of updating a video crossbar content and retrieving a video crossbar content;

a property to retrieve picture composition capabilities of the video crossbar;

15

a property to do one of setting a periodicity of an interrupt service routine and getting a periodicity of an interrupt service routine;

a property to do one of setting a time to evaluate whether a speaker is continuing to speak and getting a time to evaluate whether a speaker is continuing to speak;

20

a property to do one of setting a second time during which a speaker and a video switching process can not be taken over by a second speaker and getting a second time during which a speaker and a video switching process can not be taken over by a second speaker; and

a property to do one of setting a third time and getting a third time, the third time being the time when a switch is made and when a fast update request is sent to the speaker's system.

5 26. The apparatus according to claim 21 wherein the at least one property set comprises a decoder property set.

27. The apparatus according to claim 26 wherein the decoder property set comprises:

10 a property to specify that a video frame update be completed and a video frame be displayed until receiving a release signal; and
a property to indicate a video temporal and spatial trade-off of an encoder.

28. The apparatus according to claim 21 wherein the at least one property set comprises a video encoder send property set.

15

29. The apparatus according to claim 28 wherein the at least one hardware module comprises a video encoder, the video encoder send property set comprises:

a property to signal to the application that it needs to send a command to the video encoder.

20

30. The apparatus according to claim 21 wherein the at least one property set comprises a stream topology property set.

31. The apparatus according to claim 30 wherein the stream topology property set comprises:

a property to retrieve a direction and crossbar positional index of a stream.

5 32. The apparatus according to claim 21 wherein the at least one property set comprises a video encoder property/set.

33. The apparatus according to claim 32 wherein the video encoder property set comprises:

10 a property to command a video output stream to enter a fast update picture mode;

a property to command the video output stream to perform a fast update of a group of blocks;

15 a property to command the video output stream to perform a fast update of a macroblock;

a property to command the video output stream to use sync for every group of blocks; and

a property to provide an indication that a set of macroblocks has been received with errors and has been treated as not coded.

20

34. The apparatus according to claim 21 wherein the at least one property set comprises a network statistics property set.

35. The apparatus according to claim 32 wherein the network statistics property set comprises:

5 a property to do one of informing a video output pin of error channel conditions and supplying a media service provider component the error channel conditions; and

a property to do one of informing the video output pin of a channel packet rate loss and supplying the media service provider component the channel packet rate loss.

10 36. The apparatus according to claim 21 wherein the at least one property set comprises a bandwidth property set.

37. The apparatus according to claim 36 wherein the bandwidth property set comprises:

15 a property to do one of specifying an upper limit in bandwidth transmission to a video output pin and supplying the upper limit bandwidth transmission of the video output pin to a media service provider component.

38. The apparatus according to claim 21 wherein the at least one property set comprises a frame rate property set.

20 39. The apparatus according to claim 38 wherein the frame rate property set comprises:

a property to do one of specifying a video frame's average display time to a video output pin and supplying the video frame average display time to a media service provider component.

5 40. The apparatus according to claim 21 wherein the at least one property set comprises a RTP control property set.

41. The apparatus according to claim 40 wherein the RTP control property set comprises:
a property to do one of retrieving a maximum RTP packet size and setting
10 the maximum RTP packet size.

42. A computer-readable medium having computer-executable instructions for bridging a plurality of multicast conferences, each of the plurality of multicast conferences having at least one client, the computer-executable instructions performing the steps
15 of:

receiving a first call from one of the at least one client to join a conference;

looking for the conference; and

joining the one of the at least one client into the conference, the step of

20 joining comprising:

creating a second call to call the conference;

creating at least one multicast bridging terminal;

selecting one of at least one audio stream and at least one video stream onto the at least one multicast bridging terminal;
connecting the second call; and
answering the first call.

5

43. The computer-readable medium of claim 42 wherein the at least one multicast bridging terminal comprises one of at least one audio bridge terminal and at least one video bridge terminal.

10

44. The computer-readable medium of claim 43 wherein the at least one multicast bridging terminal comprises:

A sink module to receive at least one input stream from one of the first call and one of the second call;

a source module to send at least one output stream to one of the first call

15

and one of the second call; and

an interface to send one of at least one input stream to the source module.

20

45. The computer-readable medium of claim 44 wherein a data format of the at least one input stream and a data format of the at least one output stream is identical

46. The computer-readable medium of claim 45 wherein the at least one input stream is an audio stream and the at least one output stream is an audio stream, the data format being PCM linear at 16 bits per sample at 8 KHz.

47. The computer-readable medium of claim 45 wherein the at least one input stream is a video stream and the at least one output stream is a videostream, the data format being RTP H.263.

5

48. The computer-readable medium of claim 44 wherein the sink filter uses a memory allocator in an output pin of an upstream module, the upstream module sending the at least one input stream to the sink filter.

10 49. The computer-readable medium of claim 44 wherein the sink module is an audio sink module and the at least one input stream is at least one input audio stream, the computer-executable instructions further comprising the step of timestamping, by the audio sink module, audio samples in the at least one audio input stream with a time of a clock of the audio sink module.

15

50. The computer-readable medium of claim 49 further comprising the step of updating the clock when a discontinuity flag is set.

51. The computer-readable medium of claim 50 wherein the discontinuity flag is set
20 when a first sample of a talk spurt is delivered to the audio sink filter.

52. The computer-readable medium of claim 50 further comprising the steps of:

if the data in the at least one input stream is continuous data, increasing the clock by a first time, the first time based on an amount of data passed through the audio sink module; and

5 if there is a silence period in the at least one audio input stream, adjusting the clock by a second time, the second time being the length of time of the silence period.

10 53. The computer-readable medium of claim 45 wherein the data in the input stream is in frames of a first size and the data in the output stream is in frames of a second size, the computer-executable instructions further comprising the steps of:

calling, by the sink module, the interface to send data samples of the first size to the source filter;

if the first size is equal to the second size, sending the data in the input stream directly down stream; and

15 if the first size is not equal to the second size, constructing, by the source module, new data frames of the second size, transforming the data samples of the first size into data samples of the second size, copying the data samples of the second size into the new data frames, and sending the new data frames down stream.

20

54. The computer-readable medium of claim 53 wherein the sink module is a video sink module, the at least one input stream is at least one input video stream, the video data in the at least one input video stream is in video frames, the video frames containing

resume sending video data down stream.